Article

Designing Research With Qualitative Comparative Analysis (QCA): Approaches, Challenges, and Tools

Eva Thomann[1] and Martino Maggetti[2]

Abstract

Recent years have witnessed a host of innovations for conducting research with qualitative comparative analysis (QCA). Concurrently, important issues surrounding its uses have been highlighted. In this article, we seek to help users design QCA studies. We argue that establishing inference with QCA involves three intertwined design components: first, clarifying the question of external validity; second, ensuring internal validity; and third, explicitly adopting a specific mode of reasoning. We identify several emerging approaches to QCA rather than just one. Some approaches emphasize case knowledge, while others are condition oriented. Approaches emphasize either substantively interpretable or redundancy-free explanations, and some designs apply an inductive/explorative mode of reasoning, while others integrate deductive elements. Based on extant literature, we discuss issues surrounding inference with QCA and the tools available under different approaches to address these issues. We specify

[1] Department of Politics, University of Exeter, United Kingdom, Institute of Political Science, Heidelberg University, Heidelberg, Germany
[2] Institute of Political, Historical and International Studies (IEPHI), University of Lausanne, Lausanne, Switzerland

Corresponding Author:
Eva Thomann, Department of Politics, University of Exeter, Amory Building, Rennes Drive, EX4 4RJ Exeter, United Kingdom.
Email: e.thomann@exeter.ac.uk

trade-offs and the importance of doing justice to the nature and goals of QCA in a specific research context.

Keywords

> QCA is not just another (computer-based) data analysis technique. In order to do justice to its underlying epistemology, it needs also to be understood—and applied—as a research approach.
>
> Schneider and Wagemann (2010:398)

Since Charles Ragin launched "The Comparative Method" in 1987, the methodology and the use of qualitative comparative analysis (QCA) in its different variants have developed impressively (Rihoux and Marx 2013). QCA applications have spread across various disciplines (Rihoux et al. 2013) and have expanded from small- and intermediate-N applications to large-N studies (Fiss, Sharapov, and Cronqvist 2013; Greckhamer, Misangyi, and Fiss 2013). Concurrently, issues surrounding the use of QCA have been highlighted, targeting either QCA as a method (e.g., Hug 2013; Krogslund, Choi, and Poertner 2015; Lucas and Szatrowski 2014; Paine 2015; Seawright 2014; Tanner 2014) or current practices (e.g., Baumgartner 2015; Baumgartner and Thiem 2017, 2020; Cooper and Glaesser 2016a, 2015c; Thiem 2017, 2016b; Thiem, Baumgartner, and Bol 2016; Wagemann, Buche, and Siewert 2016). Lastly, a host of methodological advancements have taken place, not all of which have made their way into broader QCA practice (e.g., Baumgartner 2015; Baumgartner and Thiem 2017; Braumoeller 2015; Eliason and Stryker 2009; Garcia-Castro and Ariño 2016; Haesebrouck 2015; Maggetti and Levi-Faur 2013; Marx and Duşa 2011; Mikkelsen 2015; Ragin and Schneider 2011; Rohlfing and Schneider 2013; Schneider and Rohlfing 2013, 2016; Schneider and Wagemann 2013; Skaaning 2011; Thiem 2014b; Thiem and Duşa 2013; Thiem, Spöhel, and Duşa 2016).

To the average QCA user, this profusion can be perplexing. Is QCA a valid method for empirical research? What are the quality criteria? These questions have no ready-made answers, and practices vary considerably. The relatively young QCA methodology has not yet reached a high level of standardization. Indeed, QCA methodologists themselves disagree on several aspects of the method. Distinguishing QCA as a technique from QCA as an

approach helps to understand these controversies. The former denotes formalized data analysis based on data set observations, involving truth table analysis and logical minimization (Rihoux and Ragin 2009). Instead, QCA as an encompassing approach refers to research design issues, specifically "the processes before and after the analysis of the data, such as the (re-)collection of data, (re-)definition of the case selection criteria, or (re-)specification of concepts" (Schneider and Wagemann 2012:11). Many recent disputes "incorrectly infer flaws in the principles of QCA from problems in its current practice in empirical research" (Rohlfing and Schneider 2014:28). We believe that rather than focusing on QCA as a technique, such issues are better understood and disentangled by also considering QCA as an approach.

In this article, we seek to help users design QCA studies. To this end, we review, systematize, and clarify different approaches, challenges, and tools surrounding the design of research applying QCA. We provide a conceptual map with which to situate a given QCA study within the targeted approach and play by the corresponding rules. Additionally, we provide an overview of relevant literature to consult for a more in-depth treatment of these issues. Our goal is neither to address criticisms of QCA (see, e.g., Rohlfing and Schneider 2014; Thiem 2014b; Thiem, Baumgartner, et al. 2016) nor to prescribe specific standards to execute the analysis (see, e.g., Schneider and Wagemann 2010, 2012; Wagemann and Schneider 2015). Rather, our goals are to show that researchers are confronted with a variety of research strategies and tools for conducting QCA, explain how this can be done coherently in line with specific research goals, and pinpoint areas in which these approaches are incompatible. In this respect, we put forward two main arguments.

Firstly, like most empirical social research, QCA studies typically involve learning about facts we do not know by using the facts we do know—that is, they establish inference. Extant literature tells us that establishing inference entails addressing three intertwined, main components coherently: first, clarifying the question of external validity; second, establishing measurement and, more generally, internal validity; and third, adopting a mode of reasoning (Adcock and Collier 2001; Blatter and Blume 2008; Brady and Collier 2010; Cook and Campbell 1979; King, Keohane, and Verba 1994; Maggetti, Radaelli, and Gilardi 2012; Mahoney and Goertz 2006). The QCA approach is traditionally considered to tackle these issues through a strong case orientation (Rihoux 2013). This entails the analysis of a small- or intermediate-sized set of purposively selected cases that allow for modest generalization (external validity); the use of in-depth case knowledge to ensure internal validity; and an inductive, explorative,

and iterative mode of reasoning (Ragin 1987, 2000; Rihoux and Ragin 2009; Schneider and Wagemann 2012).

Although QCA as a technique fits such an approach particularly well, several alternative approaches to QCA have emerged. The QCA approach is inherently multimethod (Berg-Schlosser et al. 2009; Ragin 2000, 2008b). Since it entails a back-and-forth between ideas and evidence, it can integrate "qualitative" or "quantitative" components in different ways. Inferences with QCA can be complemented through its combination with qualitative within-case studies and process tracing (Beach and Rohlfing 2015; Mikkelsen 2015; Rohlfing and Schneider 2013; Schneider and Rohlfing 2013, 2016) as well as with statistical techniques (Cooper and Glaesser 2016b; Eliason and Stryker 2009; Fiss et al. 2013; Greckhamer et al. 2013). Indeed, QCA is increasingly being applied to large samples, typically without a qualitative element. Furthermore, contrasting procedures for deriving internally valid QCA results have been proposed (Baumgartner 2015; Baumgartner and Thiem 2017; Schneider and Wagemann 2012, 2013). Moreover, a growing number of empirical applications evaluate set-theoretic hypotheses (Ragin 1987; Schneider and Wagemann 2012; Thomann 2015). Hence, the time seems ripe to take stock of different emerging approaches to QCA and issues surrounding them.

In this vein, we identify different approaches to QCA. Some approaches are intensively case oriented, while others are condition oriented. Approaches emphasize either the substantive interpretability of results or the absence of redundancies; some designs apply an inductive and/or explorative mode of reasoning, while others integrate deductive elements. It is crucial to understand which tools are useful for which components of inference, depending on the chosen approach. To this end, we map the currently available tools for researchers adopting these approaches to reach inference in light of recently raised research design issues.

In the next section, we briefly introduce core notions about QCA and inference. We then propose our three-tiered typology of approaches to QCA. We illustrate how these approaches articulate the three components of inference and discuss the core issues and available tools. Our conclusion highlights the importance of doing justice to the nature and goals of QCA approaches in any specific research context.

Designing Research With QCA

Regardless of the chosen approach, the QCA technique explores the presence of logical implications or set relations in terms of necessity and sufficiency.

Deterministically speaking, a condition X is necessary ($\leftarrow$) for an outcome Y if X is also given whenever Y is given (i.e., Y implies X; Y is a subset of X). X is sufficient ($\rightarrow$) for Y if Y also occurs whenever X occurs (i.e., X implies Y; also, X is a subset of Y). QCA can model quasi-necessity or quasi-sufficiency more probabilistically by integrating parameters of fit and/or degrees of set membership, where the presence of a subset changes the degree of, but does not assure, the presence of the superset. QCA then models three aspects of causal complexity (Berg-Schlosser et al. 2009). This includes the assumption of causal asymmetry: The conditions explaining the occurrence of an outcome can differ from those explaining its nonoccurrence. Furthermore, "the assumption of equifinality allows for different, mutually non-exclusive explanations of the same phenomenon. Instead of assuming isolated effects of single variables, the assumption of conjunctural causation foresees the effect of a single condition unfolding only in combination with other (. . .) conditions" (Schneider and Wagemann 2012:78, emphasis removed). Since truth table analysis treats any configuration as equally important regardless of its empirical relevance, QCA reacts very sensitively to the presence or absence of singular cases (Skaaning 2011).

It is important to take seriously the fact that QCA cannot be applied to just any type of research question (Schneider and Wagemann 2010:399; Thiem, Baumgartner, et al. 2016). QCA has different uses, both descriptive and explanatory, which include summarizing data, creating typologies, checking the coherence of subset relations, evaluating existing hypotheses, testing atheoretical conjectures, and developing new theories (Berg-Schlosser et al. 2009; Schneider and Wagemann 2010). In its explanatory uses, on which this article focuses, QCA can address research questions that center on (quasi-)necessity and/or (quasi-)sufficiency, entail an interest in the causes of a given effect, and include one or more aspects of complex causation. Hereafter, we assume that the reader is familiar with the basic workings of QCA analysis (see, e.g., Ragin 2008b; Rihoux and Ragin 2009; Schneider and Wagemann 2012). Assessing such research questions with QCA is aimed at establishing valid inference. Descriptive inference involves using observations to learn about other unobserved facts. Causal inference means learning about causal effects from the observed data (King et al. 1994:8).[1] Achieving valid inferences involves three main components that QCA studies should address coherently (see Figure 1).

First, inference from an analyzed set of cases, hereafter called a "sample," to a population involves a clarification of external validity—that is, the generalizability of inferences to a broader universe of relevant cases for the research question (Cook and Campbell 1979:70-80). The potential for this
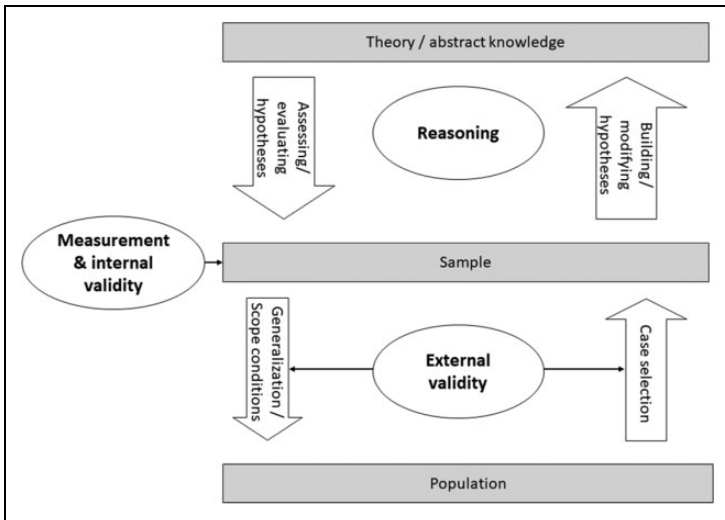
Figure 1. Components of inference. *Source:* Authors' own illustration.

extrapolation inherently depends on the case selection criteria. The quest to generate inferences that are as widely applicable as possible typically implies maximizing statistical generalization to the broader population, based on large samples that are representative (i.e., that display or mimic the distribution of characteristics in the population), and by adopting a probabilistic stance. Alternatively, comparative case studies typically select cases purposively according to theoretical criteria that determine the cases' relevance to the research question. These very criteria also constitute the scope conditions for the results—that is, the specific, explicitly defined empirical contexts within which the insights gained are deemed valid (Goertz and Mahoney 2006). This type of generalization is also called "limited," "historical," or "contingent" (Berg-Schlosser et al. 2009; Blatter and Blume 2008; Rihoux and Ragin 2009).

Second, the descriptive inferences drawn within the sample need to achieve measurement validity, which refers to whether the observations meaningfully capture the ideas contained in the concepts (Adcock and Collier 2001). Internal validity then means that the drawn inferences adequately represent the cases being studied (Brady and Collier 2010; Cook and Campbell 1979:70-80).

Third, inference typically involves making a connection between the data and reasoning (King et al. 1994). Research that aims to assess hypotheses
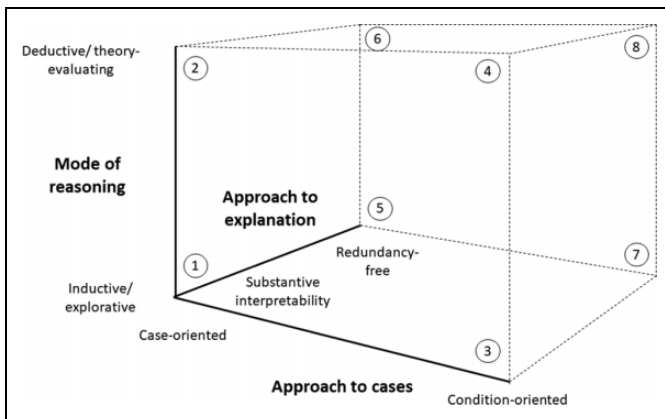
Figure 2. Approaches to QCA. _Source:_ Authors' own illustration.

starts out with a hypothetical statement, usually derived from existing knowledge (e.g., theory) or developed ad hoc, which is then compared against, supported, or refuted by empirical observations. Conversely, some research designs build or modify hypotheses after the empirical analysis. They start with the data analysis and derive specific conclusions or broader theoretical statements from it (Maggetti et al. 2012; Rohlfing 2012:9).

While distinct, these three components of inference are closely intertwined. For example, inherent trade-offs exist between the depth and accuracy of explanation (internal validity) and the explanation's empirical breadth (external validity). As we outline below, different approaches to QCA address these three components and the associated trade-offs in divergent ways.

Approaches to QCA

We argue that the current approaches to QCA differ on three main axes: in their approach to cases, their approach to explanation, and their mode of reasoning (see Figure 2).

Perhaps the most prominently discussed differentiation concerns the approach to cases. All QCA studies are configuration oriented, since they conceive of cases as a configuration of attributes (Rihoux 2013:238). However, QCA approaches differ in the emphasis that researchers put on cases as an object of inquiry (Greckhamer et al. 2013). The focus of the traditional _case-oriented_ approach is the close analysis of particular cases using deep

contextual knowledge. Here, in addition to cross-case inference, in-depth case knowledge plays a pivotal role in establishing measurement and internal validity. Such knowledge emerges from an intensive qualitative engagement with the cases often based on purposively selected small- to intermediate-N samples (Berg-Schlosser and De Meur 2009). However, the QCA technique can be applied to different sample sizes, and case numbers alone do not justify its use (Schneider and Wagemann 2012:12; Thiem 2017). *Condition-oriented* applications understand cases primarily in terms of a well-defined set of conditions. The results are mainly interpreted as patterns across cases and are not complemented with an in-depth, qualitative treatment of individual cases (Greckhamer et al. 2013; Seawright and Collier 2010:358). This approach typically uses QCA on large samples, which are often implicitly or explicitly deemed representative of an underlying population. This largely precludes intimacy with all cases but facilitates resorting to complementary statistical techniques and parameters to evaluate QCA models (Cooper and Glaesser 2016b; Fiss et al. 2013; Greckhamer et al. 2013).

Rather than the sheer number of observations, the relative closeness or distance to empirical cases distinguishes these two approaches. Whereas the case-oriented approach emphasizes the complementary use of within-case knowledge, the condition-oriented approach predominantly relies on cross-case inference, focusing on relations between sets and based on knowledge of conceptual relationships rather than of particular cases. This orientation toward conditions is sometimes found in relatively small-N analyses, while a large N does not preclude an interest in particular cases. Namely, set-theoretic multimethod research provides case selection criteria for assisting inferences derived from cross-case comparisons with targeted within-case studies (Beach and Rohlfing 2015; Mikkelsen 2015; Rohlfing and Schneider 2013; Schneider and Rohlfing 2013, 2016).

A second distinction concerns the approach to explanation, as it is usually understood in observational studies in the social sciences. One approach emphasizes the *substantive interpretability* of QCA results from a research-practical perspective, within which social research "is built upon a foundation of substantive and theoretical knowledge, not just methodological technique" (Ragin 2008b:173). Hence, the purpose of QCA is "to find meaningful super- and/or subsets of the phenomenon to be explained" (Schneider 2016:2). When analyzing sufficient conditions, this approach assesses the plausibility of counterfactual assumptions. According to this approach, the parsimonious solution assumes all logical remainders that help eliminate redundancies to be sufficient for the outcome, irrespective of the "goodness" of the counterfactual. Avoiding this can either entail deriving a

conservative (or complex) solution that assumes that empirically unobserved configurations (logical remainders) are insufficient for the outcome or an intermediate solution based on carefully justified counterfactual arguments (Ragin 2008b; Schneider and Wagemann 2012, 2013).

Additionally, this approach interprets selected necessary conditions as crucial explanatory factors, without which a given event could not have occurred (Goertz 2006; Goertz and Starr 2003; Schneider and Wagemann 2012): "In addition to empirical support (. . .), there must be theoretical and conceptual arguments as to why it is plausible to declare a given condition as necessary for an outcome" (Schneider and Wagemann 2016:317). The empirical importance of necessary and sufficient conditions is assessed in a second analytic step. Necessary conditions become empirically more important as they also approximate a sufficient condition, and sufficient conditions become so as they approximate a necessary condition (Goertz 2006; Mahoney and Sweet Vanderpoel 2015:70).

Another approach emphasizes *redundancy-free models* by highlighting that "the crucial mechanism of QCA that turns necessary and sufficient conditions into causally interpretable necessary and sufficient conditions is the elimination of redundancies" (Thiem and Baumgartner 2016a:3). While a host of supersets or subsets of an outcome exists, this approach only derives causality from conditions that are *both* minimally sufficient *and* contained in a minimally necessary condition for an outcome. That is, only a parsimonious solution that effectively eliminates all causally irrelevant (redundant) factors and has very high coverage (indicating necessity) is causally interpretable (Baumgartner 2015; Baumgartner and Thiem 2017; Thiem and Baumgartner 2016b). This approach dissociates the concept of necessity without sufficiency from that of causality. Every superset of a certain instance is necessary for that instance, but this does not mean that it is causally interpretable (Thiem 2017, 2016b; Thiem, Baumgartner, et al. 2016). Therefore, this approach considers it meaningless to propose criteria for the causal interpretation of necessary or sufficient conditions that are identified prior to minimization and are not redundancy-free (Thiem and Baumgartner 2016a:4).

These two approaches present contrasting strategies to maintaining internal validity in light of "noisy" social science data that entail the so-called limited diversity and/or less-than-perfect subset relations (Schneider and Wagemann 2012:119). They also diverge regarding the existence of criteria that render (certain) necessary and sufficient conditions causally interpretable. Yet, both approaches refer to the INUS[2] theory of causation (Ragin 2000, 2008b; Schneider and Wagemann 2012; Thiem, Baumgartner, et al. 2016; Thiem and Baumgartner 2016a). They also agree that "a set relation alone is

not enough to postulate a cause" (Schneider 2016:2; Thiem, Baumgartner, et al. 2016). The use of the QCA algorithm only describes the consequences attributable to some treatment but does not achieve a full explanation—that is, the clarification of the mechanisms through which that (potentially) causal relationship holds (Cook and Campbell 1979:7).[3] In this respect, the proponents of the approach emphasizing substantive interpretability argue that the multimethod character of the QCA approach, if applied coherently, crucially helps to complement inferences (Beach 2017; Rihoux and Ragin 2009; Schneider and Wagemann 2010:400).

Finally, QCA studies differ in their modes of reasoning. QCA is often employed to *inductively and/or exploratively* "help the researcher generate some new insights, which may then be taken as a basis for a further theoretical development or for reexamination of existing theories" (Berg-Schlosser et al. 2009:16). We define as inductive an approach whose primary aim is to build or modify a hypothesis after the analysis or to generate a new theory emerging from the empirical evidence (Rohlfing 2012). However, QCA applications increasingly explicitly formulate a priori expectations against which they compare their results. The primary aim of what we define as a *deductive* approach to QCA is to evaluate existing knowledge rather than generate new knowledge.

While these two approaches adopt different modes of reasoning, they do not correspond to the ideal–typical notions of inductive designs nor to those of deductive research designs (Eliason and Stryker 2009). First, QCA as an approach has an inherent iterative element that involves conceptual and theoretical considerations: Researchers engage in a back-and-forth between prior knowledge and cases. Theories, explanatory frameworks, concepts, and analytic decisions are refined based on preliminary empirical insights gained throughout the analysis; sampling and measurement decisions are respecified using theoretical or conceptual insights (Berg-Schlosser et al. 2009:6; Schneider and Wagemann 2012:11). Second, truth table analysis inherently entails a search for results, rather than simply testing the consistency and coverage of previously defined set-theoretic hypotheses (Thiem 2017).

The next section illustrates how the different approaches displayed in Figure 2 rely on different tools to clarify external validity, ensure internal validity, and engage in reasoning (Figure 1).

Clarifying External Validity

External validity is high when inferences about the sample under scrutiny can be generalized widely beyond its boundaries. QCA's case sensitiveness

poses challenges to external validity (e.g., Krogslund et al. 2015). Conventional statistical approaches resort to three principal remedies: first, preventing sampling bias or applying matching techniques to improve the representativeness of the sample; second, using statistical tests to indicate generalizability; and third, increasing the sample size to diminish the impact of "outliers." These strategies may neither have the same effects nor always be preferable over other options for QCA studies, depending on the adopted approach.

Specifying the Empirical Scope of the Results

Achieving external validity is particularly tricky in the absence of a representative sample. Under a *case-oriented approach*, drawing random samples is often an undesirable strategy; rather, cases are selected for which obtaining in-depth knowledge is crucial, relevant, and feasible for answering the research question (Mahoney and Goertz 2006).

As Figure 1 illustrates, the mode of reasoning of a research study (upward) is intertwined with, but does not equate to, the question of how it relates to the underlying population (downward; contrary to, e.g., Hug 2013). The deductive assessment of existing knowledge typically involves evaluating its applicability across a wide range of cases. Hence, case-oriented analysis more often proceeds inductively. However, small-N, case-oriented deductive studies can also make conclusions about the applicability of propositions to cases that satisfy the scope conditions—although this precludes an interpretation in terms of more general applicability. Case-oriented QCAs should generally define scope conditions that provide evidence about the relevant factors to explain negative and positive findings as well as delimit the context in which systematic relations or hypotheses apply (Foschi 1997; Schneider and Rohlfing 2016).

Selecting cases nonpurposively can have other analytic advantages (Seawright 2002). Statistical generalization—often sought by deductive studies—requires experimental designs or randomly drawn, representative, large, and robust samples. Yet, such techniques are also available for inductive or explorative designs (e.g., Cooper and Glaesser 2016b; Misangyi and Acharya 2014). A challenge for *condition-oriented studies* can be that inference is neither assisted by inferential statistics nor by comprehensive case intimacy. A large N alone does not ensure generalizability. Yet, if such samples are representative, then the QCA tool kit does offer statistical measures of uncertainty for necessity and sufficiency claims (Braumoeller and Goertz 2000; Eliason and Stryker 2009; Goertz, Hak, and Dul 2013; Longest

and Vaisey 2008; Ragin 2000; Seawright 2002). Most recently, Braumoeller (2015) proposed a permutation test for crisp- and fuzzy-set QCA. Few QCA applications resort to statistical inference, partly because its use on full or purposive samples yields nonsensical inference to a nonexistent population. Conversely, such tests can arguably be a useful substitute for case knowledge to indicate the degree of uncertainty of the results derived from large, representative samples.

It is striking how few contemporary QCA applications explicitly deal with question of their external validity. Specifying the empirical scope of the argument, congruent with the case selection rationale, helps to avoid confusion about the validity of different uses of QCA.

Robustness to Adding or Dropping Cases

Researchers who seek to evaluate the applicability of results in other contexts can replicate their analysis on different samples (Skaaning 2011). With a case-based method, adding cases does not necessarily "average out" the effect of "outliers" on the results, especially when new cases add new configurations to the analysis. The extent to which adding cases could substantially affect the interpretation of results can be checked through robustness tests (Maggetti and Levi-Faur 2013). *Case-oriented analyses* of small, purposive samples might be particularly sensitive to the inclusion or removal of cases—but testing for this may primarily matter to researchers who are interested in inferring beyond the sample or in the absence of unambiguous, sound case selection criteria and scope conditions.

Instead, dropping or adding cases mostly makes sense if the researchers lack close familiarity with the cases in *condition-oriented*, large-*N* analyses. In this situation, the occurrence of very rare configurations might be derived from measurement error. To reduce the latter's impact on the results, researchers can set a frequency threshold for truth table rows. Configurations below the threshold, and the cases populating them, can be ruled out (Maggetti and Levi-Faur 2013).

Scope of the Counterfactual Arguments

For scholars emphasizing *substantive interpretability*, parsimonious and intermediate solution terms can include configurations that were not empirically observed but might occur in other settings. The problem with choosing counterfactual cases may lie in "drawing too many inferences on too little information" (Schneider and Wagemann 2016:320; Wagemann et al. 2016)

or in making inferences that are difficult to interpret. Yet, avoiding this comes at the price of external validity. Conversely, the approach prioritizing *redundancy-free models* contends that relying on other solutions than the parsimonious one undermines the internal validity of results (Baumgartner and Thiem 2017). According to this approach, only the parsimonious solution—which, depending on the algorithm used, can arguably also be derived without active reliance on logical remainders—is causally interpretable (Baumgartner 2015). In sum, clarifying the external validity of QCA results involves a justification of whether assumptions about logical remainders are made, transparency regarding the implications of making assumptions on logical remainders by indicating the extent of limited diversity, and if applicable, the directional expectations and simplifying assumptions applied (Schneider and Wagemann 2010; Thomann 2015; Wagemann and Schneider 2015).

Ensuring Measurement Validity and Internal Validity

Measurement validity and internal validity require systematic bias to be removed from the analysis and the researcher to be confident that the descriptive and explanatory inferences are valid for the observations under scrutiny. Both the degree of case orientation and the approach to explanation affect how internal validity is established.

Dealing With Measurement Error

Some imprecision, uncertainty, and randomness is unavoidable in empirical analysis and stems, for example, from incomplete conceptual formation, imperfect operationalization, data problems, and the shortcomings of data analysis techniques. Because QCA offers no direct way to incorporate the error term, using an explicit procedure for dealing with measurement error would improve the validity of the results (Maggetti and Levi-Faur 2013). Hug (2013:252) claims that "scholars employing QCA rarely reflect on the possibility that the data they have gathered and used in their analysis might be error-prone and thus affect their conclusions" (see also Krogslund et al. 2015; Lucas and Szatrowski 2014). Recent applications counter that claim. Different approaches offer several tools for addressing measurement error (Cooper and Glaesser 2016b; Emmenegger, Schraff, and Walter 2014; Greckhamer et al. 2013; Maggetti and Levi-Faur 2013; Ragin 2008a; Rohlfing and Schneider 2014; Skaaning 2011; Thiem 2014a; Thiem, Spöhel, et al. 2016).

Case-oriented researchers develop an in-depth knowledge of cases and concepts, thus minimizing ex ante measurement error (Schneider and Wagemann 2012:11). When *condition-oriented* researchers lack qualitative case knowledge and/or in the absence of a priori guidance on the best model specifications, a number of complementary strategies can assist in ensuring measurement and internal validity (e.g., Cooper and Glaesser 2016b). These strategies depend on the type of error that is expected to prevail (Maggetti and Levi-Faur 2013).

First, when a systematic inaccuracy in the coding is suspected, an adjustment factor can be applied to consistency scores (see Maggetti and Levi-Faur 2013). Second, potential condition errors can be identified by comparing conservative, intermediate, and parsimonious solutions across different model specifications. This enables the detection of conditions and solution terms that are more or less robust, indicating potential measurement issues (*approach emphasizing substantive interpretability*). Ex post process tracing of deviant cases can help to discern potentially omitted conditions and measurement error (*case-oriented approach*; Rohlfing and Schneider 2013). Third, the noise created by random errors can be reduced through the use of probabilistic criteria (see above) in *condition-oriented*, large-N studies. Cooper and Glaesser (2016b) used bootstrapping for this purpose. Fourth, errors related to model specifications can be addressed through robustness tests (Skaaning 2011), by adding or dropping conditions, trying different raw consistency thresholds, and testing for different calibration strategies. The latter is particularly advisable if strong conceptual criteria are absent, especially for determining the crossover point (e.g., Sager and Thomann 2017). Emmenegger et al. (2013:190) suggest resorting to robustness tests the higher the number of cases and conditions, the less evident the choice of thresholds and the weaker the theoretical expectations. Finally, a frequency threshold allows researchers to tackle case-based errors (e.g., Misangyi and Acharya 2014).

Limited Diversity

Different approaches to QCA react differently to the possible threats to internal validity arising from the limited empirical diversity inherent in social reality. Studies using simulated data (whose adequacy is contested, especially by proponents of a case-oriented approach) suggest that limited diversity can prevent QCA from revealing the true data generating process (e.g., Krogslund, Choi and Pertner 2014; Lucas and Szatrowski 2014; Seawright 2014; see Rohlfing and Schneider 2014, for a response) or make QCA

misleadingly find "explanatory" models when faced with random data (Marx and Duşa 2011). *Case-oriented* QCA approaches may mitigate the potential problems of limited diversity for internal validity through the additional use of case knowledge. Comparative case studies can substitute for counterfactual reasoning, and single-case studies can help to support the causal interpretation of INUS conditions (Rohlfing and Schneider 2013, 2016; Schneider and Rolfing 2013, 2016). The approach emphasizing *substantive interpretability* provides criteria for the careful justification of counterfactual arguments (Ragin 2008b; Schneider and Wagemann 2012, 2013, 2015). The approach emphasizing *redundancy-free models* views limited diversity as less problematic, since parsimonious solutions reliably reveal a redundancy-free set of causal factors. However, proponents of this approach also recognize that "if the data processed by any Boolean method is deficient, parsimonious solutions will tend to miss the target just as any other type of solutions" (Baumgartner 2015:855).

Depending on the approach they adopt, QCA users probably want to reduce limited diversity a priori in their research designs (Schneider and Wagemann 2010, 2012). Thresholds for the ratio of the number of cases to the number of conditions provide heuristics to do so (Marx and Duşa 2011). Yet, what ultimately matters is the ratio of observed configurations rather than case numbers. A large *N* can, but does not have to, be a remedy. The number of conditions can be reduced, for example, through parsimonious explanatory frameworks based on cumulative knowledge (e.g., Thomann 2015); most similar system designs that hold contextual factors constant by using concept-structuring techniques (Goertz and Mahoney 2005) or two-step QCA (Schneider and Wagemann 2006).

Valid Explanation

How limited diversity is dealt with interacts with different ideas regarding a valid explanation. For scholars emphasizing *substantive interpretability*, a good explanation should be plausible and free from logical contradictions. This approach primarily views QCA results as supersets or subsets of the outcome that differ in their complexity. The parsimonious solution term consists of configurations that are either equivalent to or supersets of the configurations of the intermediate or conservative solution term. According to this approach, if the former covers a superset of (hypothetical) cases covered by the latter, then it "claims more": that is, it assumes that the outcome would have also occurred in such scenarios. Yet, the parsimonious solution term can also cover implausible or untenable configurations, and it

may therefore be less accurate than an intermediate or conservative solution term. Conversely, this approach states that if a parsimonious solution is accurate, then the intermediate and conservative solutions—as subsets of the former—by implication cannot be incorrect, although they may sometimes be overly "specific." Rather than contradicting each other, these models are supersets and subsets of each other, respectively.

To ensure the accuracy of the results, this approach entails that counterfactual claims require careful justification, regardless of the algorithm used (Emmenegger 2011). Directional expectations, based on theoretical and empirical knowledge, help to distinguish plausible (easy) from implausible (difficult) counterfactuals (standard analysis [SA]; Ragin 2008b). Simultaneous subset relations and contradictory assumptions prevail when the same configuration is incoherently considered sufficient for both the outcome and its negation or when a configuration implies the outcome despite containing a negated necessary condition (Schneider and Wagemann 2012, 2013). Such untenable and other logically impossible arguments can be avoided through appropriate treatment of remainders with enhanced SA (ESA; Schneider and Wagemann 2015; for applications, see Sager and Thomann 2017; Thomann 2015).

In contrast, the approach emphasizing *redundancy-free models* primarily views QCA results as causal claims. This approach maintains that a "configurationally correct" QCA solution only contains causally relevant factors (Baumgartner and Thiem 2017; Thiem and Baumgartner 2016b). Only the parsimonious solution can reliably reveal such Boolean difference makers. According to this approach, intermediate and conservative solution formulas cannot be causally interpreted because they still contain conditions that can be further eliminated (Baumgartner 2015:840). Hence, different degrees of complexity are more than just a matter of specificity—they are about "false positives." By attributing causal relevance to more factors than minimally needed, intermediate and conservative solutions would make *a superset* of the causal claims than the parsimonious solution (see Baumgartner and Thiem 2017, for a detailed argumentation). Hence, according to this approach, it is possible that an intermediate or conservative solution may incorrectly attribute causal relevance to some factors, whereas the parsimonious solution does not. Setting the statement of sufficiency to "false" for certain logical remainders does constitute a counterfactual assumption—one that leads researchers to make causal claims that are not actually "conservative" (Baumgartner 2015; Thiem 2016b).

It becomes clear that these two approaches adopt incompatible views on the relationship between different solution terms. Under an *approach*

emphasizing substantive interpretability, one way to enhance transparency about different solution terms is to comparatively inspect these solutions (Maggetti and Levi-Faur 2013). Researchers can identify factors of the presented solution that belong to the parsimonious solution to facilitate an informed interpretation. Fiss (2011) proposes an illustration that integrates but simultaneously distinguishes the parsimonious solution term (the "causal core") and additional factors (the "causal periphery") appearing in the intermediate or conservative solution term (see, e.g., Misangyi and Acharya 2014). While enhancing transparency, this does not overcome the contradiction for the *approach emphasizing redundancy-free models*. For *case-oriented* researchers, case knowledge can clarify causally interpretable aspects of QCA results—for example, by discussing each sufficient path through a typical case study. Set-theoretic multimethod research provides a powerful tool for replicating, confirming, or extending inferences (Beach and Rohlfing 2015; Rohlfing and Schneider 2013; Schneider and Rohlfing 2013, 2016).

Skewed Data

The large majority of cases having high or low membership in a set can seriously affect internal validity: Skewed set memberships exacerbate limited diversity and can produce simultaneous subset relations (Cooper and Glaesser 2016a; Schneider and Wagemann 2012). Different approaches grant more or less importance to these issues. In addition, skewedness strongly distorts parameters of fit, especially with fuzzy sets; for example, consistency sufficiency expresses the degree to which the statement "membership in $X \leq$ membership in Y" holds for *all* cases (Braumoeller 2017; Schneider and Wagemann 2012:127-8). Consistency sufficiency becomes artificially high as X is skewed toward zero, and it generally becomes low if X is skewed toward one. The exact opposite holds for coverage scores. These distortions tend to add up as N grows, up to the noninterpretability of the parameters of fit (cf. Cooper and Glaesser 2016b). By applying standard consistency thresholds, *condition-oriented* QCAs rarely account for the large extent to which these parameters are an artifact of case distributions. Skewedness may contribute to the often low coverage values yielded in such studies (Misangyi and Acharya 2014; Wagemann et al. 2016).

Researchers can diagnose, avoid, and report skewedness ex ante as a standard part of the research design phase by using descriptive statistics and graphical tools. Skewedness may lead researchers to iteratively respecify their sample or concepts, such as by adjusting their measurement or

calibration strategy or restructuring their concepts (Goertz and Mahoney 2005). In addition, several tools can help researchers to become aware of how their results are affected by case distributions (Cooper and Glaesser 2016a). Researchers emphasizing *substantive interpretability* can detect simultaneous subset relations using the proportional reduction in inconsistency measure. The relevance of necessity indicates the irrelevance of necessary conditions that approximate a constant (Schneider and Wagemann 2012). Schneider and Wagemann (2012:232-50) propose XY plots as a diagnostic tool. Calculating the Boolean intersections of different solution terms equally helps to detect—and ESA effectively precludes—simultaneous subset relations, which exist if the results for the positive and negative outcomes overlap. Especially for *condition-oriented* studies, updated consistency and coverage formulae could be made less sensitive to skewed set membership (Rohlfing and Schneider 2014; recently, Haesebrouck 2015).

Mode of Reasoning

Schneider and Wagemann (2012:296) observe that deductive "hypothesis testing as understood in the vast majority of applied quantitative methods does not feature among the primary goals of standard applications of set-theory based methods." Indeed, standard hypothesis testing makes the mode of reasoning (upward in Figure 1) contingent on external validity (downward in Figure 1; Ragin and Schneider 2011:150). The assertion that prior screening and adjustment of the data would distort a statistical test is at odds with the iterative elements of truth table analysis. However, Figure 1 shows that, while specifying the empirical scope of the abstract reasoning is an important component of inference, these are not equivalent analytic steps. In fact, the QCA technique can be fruitfully applied in either an exploratory, inductive research design or a confirmatory, deductive research design (Eliason and Stryker 2009; Ragin and Schneider 2011:153). Both modes of reasoning are valuable ways of contributing to knowledge and/or theory.

Building or Modifying Hypotheses

Once external validity is clarified, there is no reason why the inductive use of QCA should "fade away" (Hug 2013). As Gerring (2004:349) notes, "path-breaking research is, by definition, exploratory." In that regard, QCA can be a powerful tool to generate set-theoretic hypotheses that account for causal complexity (Ragin and Schneider 2011). *Case-oriented* studies usually generate middle range rather than grand theories (Mahoney and Goertz 2006),

whose applicability to other empirical contexts remains to be tested. An iterative, in-depth approach is a crucial strength when exploring underresearched or undertheorized phenomena, illuminating causal mechanisms, suggesting alternative theoretical explanations, and extending or refining existing knowledge (Blatter and Blume 2008; e.g., Mikkelsen 2015; Rohlfing and Schneider 2013; Schneider and Rohlfing 2013, 2016). In explorative *condition-oriented*, large-N QCA, high external validity supports more generalizable claims. Yet, iterative or inductive model specification often relies on insights gained from cases that are very demanding to obtain with a large N (Schneider and Wagemann 2010).

Hypothesis Assessment and Evaluation

Deductive approaches to QCA are especially useful when a rich body of theoretical and substantial knowledge can be assessed and refined from a set-theoretic lens (Eliason and Stryker 2009). *Condition-oriented* applications might lend themselves more to an a priori procedure of theoretically founded theory building. Yet, a small N is no obstacle to hypothesis assessment: Deterministically speaking, one deviant case suffices to falsify a set-theoretic hypothesis, especially when the research design uses "most likely" or "least likely" cases (Ragin and Schneider 2011). *Case-oriented* studies can inform about the capacity, relevance, or relative strength of theories to explain and understand the case(s). Referring to Figure 1, inference is primarily drawn "upward" in the direction of abstract knowledge rather than "downward" to the population. Hypothesis assessment with QCA typically retains an "iterative interaction between theoretical implications and empirical indications (. . .) [which] makes it possible to use the full richness of information related to the empirical case to draw inferences about the relevance of theoretical concepts" (Blatter and Blume 2008:327).

Independently of case orientation, formal theory evaluation enables researchers to systematically evaluate set-theoretic propositions against the empirical results, based on the Boolean intersections of the hypotheses, the results, and their logical negations (Ragin 1987). Going beyond traditional deductive hypothesis testing, this enables researchers to answer four questions. First, which parts of the hypothesis are supported by the findings? Second, in which directions should the hypothesis be expanded (explorative)? Third, which parts of the hypothesis should be dismissed? Fourth, which cases are the most relevant for ex post within-case analysis? Schneider and Wagemann (2012:295-304) complemented Ragin's proposal by integrating consistency and coverage to refine theory evaluation and account for how

many cases are members of the outcome and the nonoutcome in the different intersecting areas (see Sager and Thomann 2017; Thomann 2015).

Formulating Expectations in Line With QCA's Logic

The general challenge is determining how hypotheses can be meaningfully assessed with QCA (Fischer and Maggetti 2017). Expectations must be formulated in line with the logic of QCA (Schneider and Wagemann 2012). This means, first, to expect set-relational patterns of (quasi-)necessity or (quasi-)sufficiency. Hypotheses on the net correlational effects of single variables cannot be meaningfully assessed (Schneider and Wagemann 2010; Thiem, Baumgartner, et al. 2016). Directional expectations for single conditions, used for counterfactual reasoning under an approach emphasizing substantive interpretability, do *not* correspond to directly testable hypotheses. Second, aspects of complex causation can be hypothesized (e.g., equifinality or asymmetric causality, or that several factors must combine to produce an outcome, or that the effect of some factor might be contingent on the presence or absence of another factor or other factors).[4] Third, as QCA "minimizes away" irrelevant factors, expectations can target the relevance or irrelevance of some factors to an outcome.

Implicit assumptions about set-theoretic or complex causality are more widespread in social science theories than is sometimes claimed. Necessary condition hypotheses and assumptions about conjunctural causation and equifinality are actually ubiquitous (Fischer and Maggetti 2017; Goertz and Starr 2003). Set-theoretic methods are ideally suited to assess such assertions (Emmenegger et al. 2013).

Deductive and Explorative Approaches to Analyzing Necessity

The Fuzzy-Set / Qualitative Comparative Analysis (fs/QCA) software (Ragin and Davey 2014) restricts the analysis of necessity to the essentially *deductive testing* of previously defined single conditions or theoretically interesting disjunctions (Schneider and Wagemann 2010). This enables tests of predefined expectations on individual necessary conditions or disjunctions of conditions. This procedure avoids inferring necessity from sufficient conditions and acknowledges that analyzing single necessary conditions suffices to detect necessary conjunctions (Schneider and Wagemann 2012:73-74).

In the absence of clear a priori expectations, the superset/subset analysis offered by the R packages QCA (Duşa 2007) and QCApro (Thiem 2016a) facilitate a genuinely *inductive or explorative* search of all possible supersets

of the outcome—single or disjunctions of conditions (Duşa 2007; Thiem 2017, 2016a; Thiem and Duşa 2013). Superset/subset analysis typically reveals a host of supersets: the logical or makes the disjunction set larger and hence a more likely superset of the outcome (Schneider and Wagemann 2012:74).

The approach to explanation influences which of these strategies is viable and which is not. Researchers prioritizing *redundancy-free models* will never separately or deductively analyze single necessary conditions or disjunctions of single conditions because necessary conditions are not considered causally meaningful if they are not redundancy-free (disjunctions of) minimally sufficient conjunctions (Thiem 2016b). Hence, no separate search for necessity is envisaged, and no substantive interpretation of that necessity claim is deemed relevant. Conversely, researchers emphasizing *substantive interpretability* will analyze necessity separately, but they caution that supersets of the outcome can only be interpreted as meaningful necessary conditions if there are strong and plausible arguments that the conditions combined by the logical OR represent some higher-order construct (Schneider and Wagemann 2012:74). Hence, they may wish to prioritize disjunctions that are of analytical interest, operating as functional equivalents (Goertz and Mahoney 2005; Schneider and Wagemann 2010). Necessary conditions become trivial as the resulting disjunction set approximates a constant (Schneider and Wagemann 2012:235-9).

Discussion

We have illustrated different options available to QCA users to assist with inference in QCA. However, our message is not that "anything goes." Contrariwise, below we discuss the applicability of tools and associated trade-offs, depending on the corresponding approach (see Table, Online Appendix).

Compatibility Between Tools and Approaches

The use of any approach requires coherent choices throughout the analysis. Indeed, our core argument is that any given QCA study should select and justify one approach to cases, one approach to explanation, and one mode of reasoning *and remain faithful to them*. For example, adopting a genuinely explorative mode of reasoning would also indicate checking for necessary conditions that were not hypothesized a priori. Similarly, studies

emphasizing substantive interpretability explicitly discuss the empirical importance of both necessary and sufficient conditions as well as their interpretability in light of existing theoretical and substantive knowledge. Case-oriented QCAs complement their analysis with case knowledge before, during, and after truth table analysis and define the scope conditions.

Furthermore, some tools can remedy challenges to validity regardless of the approach. This particularly holds for techniques to remedy limited diversity prior to the analysis as well as many robustness tests. Other tools make more sense under one approach than under others. For example, condition-oriented studies tend to maximize external validity. Thus, probabilistic criteria might become useful, while such studies are more limited with respect to internal validity. The opposite often holds for case-oriented QCA, for which probabilistic techniques make little sense, especially in the absence of a large, representative sample. Table 1 suggests that trade-offs between tools are particularly pronounced for the two approaches to explanation. If there is limited diversity, QCA studies normally cannot simultaneously yield parsimonious solutions, as required by the approach emphasizing redundancy-free models, while also avoiding implausible counterfactual arguments, as needed under the approach emphasizing substantive interpretability. As these approaches make incompatible claims, a given QCA study has to "choose sides."

Usefulness and Practicability of Approaches

It goes beyond the scope of this article to assess the cogency of different approaches. However, our illustration shows that a strong condition orientation entails that researchers support their inferences with tools that typically rely on a relatively large number of cases. Hence, such an approach might be less appropriate for small-N studies. A strong case orientation, in turn, implies much effort needed to obtain case knowledge. In our view, set-theoretic multimethod research also opens up this possibility to large-N researchers. Furthermore, the approach emphasizing substantive interpretability entails relatively demanding procedures to deal with logical remainders and interpret the often complex results. Notwithstanding this, the approach is currently enjoying greater popularity among users—including the authors of this article—than the approach emphasizing redundancy-free models. The latter approach does not explicitly address issues involving untenable simplifying assumptions and simultaneous subset relations; rather, it frames such issues as a problem of the algorithm used. It also poses very

one-sided criteria for the causal interpretability of QCA results. Many QCA users and methodologists concur that single necessary (but not sufficient) conditions are critically important for social science theory and practice (Dul 2016:1516). Against this background, more research is needed to evaluate and assess these two approaches.

Finally, as Emmenegger et al. (2013:190) note, the establishment of "complex propositions or propositions formulated in set-theoretical terms (. . .) is the core advantage of QCA over statistical methods and also most interesting from a scientific point of view." While this is true for both inductive/explorative and deductive/theory-evaluating applications of QCA, the former have been much more widespread than the latter, partly due to the exploratory and iterative aspects of the QCA technique itself. However, recent applications show the potential of using formal theory evaluation for more deductive uses of QCA (Sager and Thomann 2017; Thomann 2015). We believe this to be a promising avenue through which to further explore the full potential of QCA to contribute to social science theories in a systematic, cumulative manner (Emmenegger et al. 2013; Fischer and Maggetti 2017; Fiss 2011; Goertz and Starr 2003; Mahoney and Sweet Vanderpoel 2015; Ragin and Schneider 2011). Note, however, that evaluating theories and hypotheses primarily makes sense when a prior body of knowledge allows for deriving expectations.

Combining Approaches

While a study usually adopts one approach to cases, one to explanation, and one mode of reasoning, different combinations of these approaches are thinkable. Yet, current QCA practice does not feature all eight possible ideal–typical combinations of approaches with the same frequency, although many "hybrids" in between can be observed. Future research should assess the (in)coherence and feasibility of such hybrids. In particular, empirical studies do not (yet) comprehensively apply the approach emphasizing redundancy-free models. Case-oriented approaches are especially suitable for explorative-inductive designs, whereas condition-oriented approaches can be more limited in this regard. Some caveats apply for other combinations. For example, strongly case-oriented researchers may not wish to strictly prioritize the absence of redundancies over case-related criteria for establishing internal validity. Redundancy-free models thus fit better with a condition-oriented approach. Future research should more fully explore the possibilities and limits of combining different approaches.

Conclusions

Schneider and Wagemann (2010) and others before them rightly emphasized that QCA is more than a technique. Accordingly, we argue that various approaches to designing QCA studies exist, which differ in their emphasis on cases, their conception of a valid explanation, and their mode of reasoning. Our message is simple and twofold. First, establishing valid inference typically entails addressing each of the three components of inference: clarifying external validity, ensuring measurement and internal validity, and adopting a specific mode of reasoning (Adcock and Collier 2001; Blatter and Blume 2008; Cook and Campbell 1979; King et al. 1994). Second, different approaches have different analytic goals and offer sometimes complementary, sometimes contradictory ways of doing so. Each analytical choice must therefore be made, first, coherently with the overarching analytical approach and, second, by following the specific issue at stake. While the specific tools we discussed are likely to advance, our typology of approaches to QCA should also help to systematize future developments. It should considerably help to reduce confusion surrounding the use of QCA if users clearly state their approaches in line with their epistemology and goals, play by the corresponding rules, and apply tools that are compatible with each other. It is up to QCA teachers, reviewers, and users to move the use of QCA forward. With this study, we hope to facilitate such improvements through the constructive illustration of different options to design research with QCA.

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Causal inferences, when narrowly defined, are difficult to draw based on observational data. We refer to the concept pragmatically, assuming that such studies usually seek to explain the occurrence of some phenomenon using a set of explanatory factors (see also Berg-Schlosser et al. 2009:3).
2. Insufficient but Nonredundant parts of Unnecessary but Sufficient conditions.
3. Accordingly, there exists a strand of qualitative comparative analysis (QCA) scholars who refrain from causal inference but use different concepts which are less demanding and pertain to the possibilistic nature of QCA.
4. Thiem, Baumgartner, et al. (2016) recommend refraining from using the term "interaction" in combination with QCA.

References

Adcock, R. and D. Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *The American Political Science Review* 95:529-46.

Baumgartner, M.2015. "Parsimony and Causality." *Quality & Quantity* 49:839-56.

Baumgartner, M. and A. Thiem. 2017. "Model Ambiguities in Configurational Comparative Research." *Sociological Methods & Research* 46:954-87. doi: 10.1177/0049124115610351.

Baumgartner, M. and A. Thiem. 2020. "Often Trusted But Never (Properly) Tested: Evaluating Qualitative Comparative Analysis." *Sociological Methods & Research* 49:279-311. doi: 10.1177/0049124117701487.

Beach, D. 2017. "Achieving Methodological Alignment When Combining QCA and Process tracing in Practice." *Sociological Methods & Research* 1-36. doi: 10.1177/0049124117701475.

Beach, D. and I. Rohlfing. 2015. "Integrating Cross-case Analyses and Process Tracing in Set-Theoretic Research Strategies and Parameters of Debate." *Sociological Methods & Research* 1-34. doi: 10.1177/0049124115613780.

Berg-Schlosser, D. and G. De Meur. 2009. "Comparative Research Design: Case and Variable Selection." Pp. 19-32 in *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*, edited by B. Rihoux and C. C. Ragin. Thousand Oaks, CA: Sage.

Berg-Schlosser, D., G. De Meur, B. Rihoux, and C. C. Ragin. 2009. "Qualitative Comparative Analysis (QCA) as an Approach." Pp. 1-18 in *Configurational*

Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques, edited by B. Rihoux and C. C. Ragin. Thousand Oaks, CA: Sage.

Blatter, J. and T. Blume. 2008. "In Search of Co-variance, Causal Mechanisms or Congruence? Towards a Plural Understanding of Case Studies." *Swiss Political Science Review* 14:315-56.

Brady, H. E. and D. Collier. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Plymouth, England: Rowman & Littlefield.

Braumoeller, B. F.2015. "Guarding against False Positives in Qualitative Comparative Analysis." *Political Analysis* 23:471-87.

Braumoeller, B. F. 2017. "Aggregation Bias and the Analysis of Necessary and Sufficient Conditions in fsQCA." *Sociological Methods & Research* 46:242-251.

Braumoeller, B. F. and G. Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44:844-58.

Cook, T. D. and D. T. Campbell. 1979. *Quasi-experimentation*. Boston, MA: Houghton Mifflin.

Cooper, B. and J. Glaesser. 2016a. "Analysing Necessity and Sufficiency with Qualitative Comparative Analysis: How do Results Vary as Case Weights Change?" *Quality & Quantity* 50:327-346.

Cooper, B. and J. Glaesser. 2016b. "Exploring the Robustness of Set Theoretic Findings from a Large N FsQCA: An Illustration from the Sociology of Education." *International Journal of Social Research Methodology* 19:445-459.

Dul, J.2016. "Identifying Single Necessary Conditions with NCA and fsQCA." *Journal of Business Research* 69:1516-23.

Duşa, A.2007. "User Manual for the QCA(GUI) Package in R." *Journal of Business Research* 60:576-86.

Eliason, S. R. and R. Stryker. 2009. "Goodness-of-fit Tests and Descriptive Measures in Fuzzy-set Analysis." *Sociological Methods & Research* 38:102-46.

Emmenegger, P.2011. "How Good Are Your Counterfactuals? Assessing Quantitative Macro-comparative Welfare State Research with Qualitative Criteria." *Journal of European Social Policy* 21:365-80.

Emmenegger, P., J. Kvist, and S. Skaaning. 2013. "Making the Most of Configurational Comparative Analysis: An Assessment of QCA Applications in Comparative Welfare-state Research." *Political Research Quarterly* 66:185-90.

Emmenegger, P., D. Schraff, and A. Walter. 2014. "QCA, the Truth Table Analysis and Large-*N* Survey Data: The Benefits of Calibration and the Importance of Robustness Tests." *COMPASS Working Paper*2014-2079. Available at: http://www.compasss.org/wpseries/EmmeneggerSchraffWalter2014.pdf

Fischer, M. and M. Maggetti. 2017. "Qualitative Comparative Analysis and the Study of Policy Processes." *Journal of Comparative Policy Analysis: Research and Practice* 19:345-61.

Fiss, P. C.2011. "Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research." *Academy of Management Journal* 54:393-420.

Fiss, P. C., D. Sharapov, and K. Cronqvist. 2013. "Opposites Attract? Opportunities and Challenges for Integrating Large-N QCA and Econometric Analysis." *Political Research Quarterly* 66:191-7.

Foschi, M.1997. "On Scope Conditions." *Small Group Research* 28:535-55.

Garcia-Castro, R. and M. A. Ariño. 2016. "A General Approach to Panel Data Set-theoretic Research" *Journal of Advances in Management Sciences & Information Systems* 2:63-76.

Gerring, J.2004. "What Is a Case Study and What Is It Good for?" *American Political Science Review* 98:341-54.

Goertz, G.2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41:88-109.

Goertz, G., T. Hak, and J. Dul. 2013. "Ceilings and Floors: Where Are There No Observations?" *Sociological Methods & Research* 42:3-40.

Goertz, G. and J. Mahoney. 2005. "Two-level Theories and Fuzzy-set Analysis." *Sociological Methods & Research* 33:497-538.

Goertz, G. and H. Starr. 2003. *Necessary Conditions: Theory, Methodology, and Applications*. New York, NY: Rowman & Littlefield.

Greckhamer, T., V. F. Misangyi, and P. C. Fiss. 2013. "The Two QCAs: From a Small-N to a Large-N Set Theoretic Approach." Pp. 49-75 in *Configurational Theory and Methods in Organizational Research (Research in the Sociology of Organizations, Volume 38)*, edited by P. C. Fiss, B. Cambré, and A. Marx. Bingley, UK: Emerald Group Publishing.

Haesebrouck, T. 2015. "Pitfalls in QCA's Consistency Measure." *Journal of Comparative Politics* 2:65-80.

Hug, S.2013. "Qualitative Comparative Analysis: How Inductive Use and Measurement Error Lead to Problematic Inference." *Political Analysis* 21:252-65.

King, G., R. O. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Krogslund, C., D. D. Choi, and M. Poertner. 2014. "Fuzzy Sets on Shaky Ground: Parameter Sensitivity and Confirmation Bias in FsQCA." *Political Analysis* 23: 21-41.

Longest, K. C. and S. Vaisey. 2008. "Fuzzy: A Program for Performing Qualitative Comparative Analyses (QCA) in Stata." *Stata Journal* 8:79-104.

Lucas, S. R. and A. Szatrowski, A. 2014. "Qualitative Comparative Analysis in Critical Perspective." *Sociological Methodology* 44:1-79.

Maggetti, M. and D. Levi-Faur. 2013. "Dealing with Errors in QCA." *Political Research Quarterly* 66:198-204.

Maggetti, M., C. M. Radaelli, and F. Gilardi. 2012. *Designing Research in the Social Sciences*. London, UK: Sage.

Mahoney, J. and G. Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14:227-49.

Mahoney, J. and R. Sweet Vanderpoel. 2015. "Set Diagrams and Qualitative Research." *Comparative Political Studies* 48:65-100.

Marx, A. and A. Duşa. 2011. "Crisp-Set Qualitative Comparative Analysis (csQCA): Contradictions and Consistency Benchmarks for Model Specification." *Methodological Innovations Online* 6:103-48.

Mikkelsen, K. S.2015. "Negative Case Selection: Justifications and Consequences for Set-Theoretic MMR." *Sociological Methods & Research* 1-33. doi: 10.1177/0049124115591015.

Misangyi, V. F. and A. G. Acharya. 2014. "Substitutes or Complements? A Configurational Examination of Corporate Governance Mechanisms." *Academy of Management Journal* 57:1681-705.

Paine, J.2015. "Set-Theoretic Comparative Methods: Less Distinctive Than Claimed." *Comparative Political Studies* 49: 793-800. doi: 10.1177/0010414014564851.

Ragin, C. C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press

Ragin, C. C.2000. *Fuzzy-set Social Science*. Chicago, IL: University of Chicago Press.

Ragin, C. C.2008a. "Measurement versus Calibration: A Set-theoretic Approach." Pp. 174-98 in *The Oxford Handbook of Political Methodology*, edited by J. M. Box-Steffensmeier, H. E. Brady, and D. Collier. Oxford, UK: Oxford University Press.

Ragin, C. C.2008b. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.

Ragin, C. C and S. Davey. 2014. *fs/QCA [Computer Programme]*. Version [2.5/3.0]. Irvine: University of California.

Ragin, C. C. and G. A. Schneider. 2011. "Case-oriented Theory Building and Theory Testing." Pp. 150-66 in *The Sage Handbook of Innovation in Social Research Methods*, edited by M. Williams and W. P. Vogt. Los Angeles, CA: Sage.

Rihoux, B.2013. "Qualitative Comparative Analysis (QCA), Anno 2013: Reframing the Comparative Method's Seminal Statements." *Swiss Political Science Review* 19:233-45.

Rihoux, B., P. Alamos, D. Bol, A. Marx, and I. Rezsohazy. 2013. "From Niche to Mainstream Method? A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011." *Political Research Quarterly* 66:175-84.

Rihoux, B. and A. Marx. 2013. "QCA 25 Years after "The Comparative Method": Mapping, Challenges, and Innovations—Mini-symposium." *Political Research Quarterly* 66:167-235.

Rihoux, B. and C. C. Ragin. 2009. *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks, CA: Sage.

Rohlfing, I. 2012. *Case Studies and Causal Inference: An Integrative Framework*. Basingtoke, UK: Palgrave Macmillan.

Rohlfing, I. and C. Q. Schneider. 2013. "Improving Research on Necessary Conditions: Formalized Case Selection for Process Tracing after QCA." *Political Research Quarterly* 66:220-35.

Rohlfing, I. and C. Q. Schneider. 2014. "Clarifying Misunderstandings, Moving Forward: Towards Standards and Tools for Set-theoretic Methods." *Qualitative & Multi-Method Research* 12:27-34.

Rohlfing, I. and C. Q. Schneider. 2016. "A Unifying Framework for Causal Analysis in Set-theoretic Multimethod Research." *Sociological Methods & Research* 1-27. doi: 10.1177/0049124115626170.

Sager, F. and E. Thomann. 2017. "Multiple Streams in Member State Implementation: Politics, Problem Construction and Policy Paths in Swiss Asylum Policy." *Journal of Public Policy* 37:287-314.

Schneider, C. Q. 2016. "Real Differences and Overlooked Similarities: Set-methods in Comparative Perspective." *Comparative Political Studies* 49:781-792.

Schneider, C. Q. and I. Rohlfing. 2013. "Combining QCA and Process Tracing in Set-theoretic Multi-method Research." *Sociological Methods & Research* 42:559-97.

Schneider, C. Q. and I. Rohlfing. 2016. "Case Studies Nested in Fuzzy-set QCA on Sufficiency: Formalizing Case Selection and Causal Inference." *Sociological Methods & Research* 45:526-68.

Schneider, C. Q. and C. Wagemann. 2006. "Reducing Complexity in Qualitative Comparative Analysis (QCA): Remote and Proximate Factors and the Consolidation of Democracy." *European Journal of Political Research* 45:751-86.

Schneider, C. Q. and C. Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy Sets." *Comparative Sociology* 9:397-418.

Schneider, C. Q. and C. Wagemann. 2012. *Set-theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis*. New York, NY: Cambridge University Press.

Schneider, C. Q. and C. Wagemann. 2013. "Doing Justice to Logical Remainders in QCA: Moving Beyond the Standard Analysis." *Political Research Quarterly* 66:211-20.

Schneider, C. Q. and C. Wagemann. 2016. "Assessing ESA on What It Is Designed for: A Reply to Cooper and Glaesser." *Field Methods* 28:316-321.

Seawright, J.2002. "Testing for Necessary and/or Sufficient Causation: Which Cases Are Relevant?" *Political Analysis* 10:178-93.

Seawright, J.2014. "Comment: Limited Diversity and the Unreliability of QCA." *Sociological Methodology*. 44:118-21.

Seawright, J. and D. Collier. 2010. "Glossary." Pp. 313-60 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by H. E. Brady and D. Collier. Plymouth, UK: Rowman & Littlefield.

Skaaning, S.2011. "Assessing the Robustness of Crisp-set and Fuzzy-set QCA Results." *Sociological Methods & Research* 40:391-408.

Tanner, S.2014. "QCA and Causal Inference: A Poor Match for Public Policy Research." *Qualitative & Multi-Method Research* 12:15-24.

Thiem, A.2014a. "Membership Function Sensitivity of Descriptive Statistics in Fuzzy-set Relations." *International Journal of Social Research Methodology* 17:625-42.

Thiem, A.2014b. "Unifying Configurational Comparative Methods: Generalized-set Qualitative Comparative Analysis." *Sociological Methods & Research* 43:313-37.

Thiem, A. 2016a. QCApro: Professional Functionality for Performing and Evaluating Qualitative Comparative Analysis. R Package Version 1.1-1. Accessed September 13, 2017 (http://www.alrik-thiem.net/software/).

Thiem, A.2016b. "Standards of Good Practice and the Methodology of Necessary Conditions in Qualitative Comparative Analysis." *Political Analysis* 24:478-84.

Thiem, A. 2017. "Conducting Configurational Comparative Research with Qualitative Comparative Analysis A Hands-on Tutorial for Applied Evaluation Scholars and Practitioners." *American Journal of Evaluation* 38:420-33.

Thiem, A. and M. Baumgartner. 2016a. "Back to Square One A Reply to Munck, Paine, and Schneider." *Comparative Political Studies* 49:801-806.

Thiem, A. and M. Baumgartner. 2016b. "Modeling Causal Irrelevance in Evaluations of Configurational Comparative Methods." *Sociological Methodology* 46: 345-357.

Thiem, A., M. Baumgartner, and D. Bol. 2016. "Still Lost in Translation! A Correction of Three Misunderstandings between Configurational Comparativists and Regressional Analysts." *Comparative Political Studies* 49:742-774.

Thiem, A. and A. Duşa. 2013. "QCA: A Package for Qualitative Comparative Analysis." *The R Journal* 5:1-11.

Thiem, A., R. Spöhel, and A. Duşa. 2016. "Enhancing Sensitivity Diagnostics for Qualitative Comparative Analysis: A Combinatorial Approach." *Political Analysis* 24:104-120.

Thomann, E.2016. "Customizing Europe: Transposition as Bottom-up Implementation." *Journal of European Public Policy* 22:1368-87.

Wagemann, C., J. Buche, and M. B. Siewert. 2016. "QCA and Business Research: Work in Progress or a Consolidated Agenda?" *Journal of Business Research* 69: 2531-40.

Wagemann, C. and C. Q. Schneider. 2015. "Transparency Standards in Qualitative Comparative Analysis." *Qualitative & Multi-Method Research* 13:38-42.

Author Biographies

Eva Thomann is a senior lecturer in public policy at the Department of Politics of the University of Exeter, United Kingdom. She specializes in Public Policy and Public Administration, with a focus on European Union and policy implementation research. Her QCA research appeared in the *European Journal of Political Research*, *Governance*, the *Journal of European Public Policy*, the *Journal of Public Policy*, and *Public Administration*.

Martino Maggetti is an associate professor in political science at the Institute of Political, Historical and International Studies (IEPHI) of the University of Lausanne, Switzerland. His research interests mainly focus on public policy and regulatory governance. His QCA-oriented research articles have appeared in several journals, including *Business & Society*, *Journal of Comparative Policy Analysis*, *Journal of European Public Policy*, *Journal of Public Policy*, *Political Research Quarterly*, and *Regulation & Governance*. His latest book is *Comparative Politics: Theoretical and Methodological Challenges* (Edward Elgar 2015, co-edited with Dietmar Braun).